

Conference Abstract

‘As Open as Possible, as Closed as Necessary’ - Managing legal and owner-defined restrictions to openness of biodiversity data

Kari M Lahti[‡], Leif Schulman[‡], Esko Piirainen[‡], Ville-Matti Riihikoski[‡], Aino Juslén[‡]

[‡] Finnish Museum of Natural History LUOMUS, Helsinki, Finland

Corresponding author: Kari M Lahti (kari.lahti@helsinki.fi)

Received: 17 Jun 2019 | Published: 26 Jun 2019

Citation: Lahti K, Schulman L, Piirainen E, Riihikoski V, Juslén A (2019) ‘As Open as Possible, as Closed as Necessary’ – Managing legal and owner-defined restrictions to openness of biodiversity data. Biodiversity Information Science and Standards 3: e37395. <https://doi.org/10.3897/biss.3.37395>

Abstract

The Finnish Biodiversity Information Facility FinBIF receives, stores and manages biodiversity data mobilised in Finland, and shares the data through its own portal (species.fi) and through Global Biodiversity Information Facility GBIF. FinBIF’s data policy ([data policy](#)) embraces the European FAIR data principles (FAIR - Findable, Accessible, Interoperable, Reusable; Wilkinson (2016)) but also incorporates specific restrictions stemming from national legislation, researchers’ needs, and data owners’ requirements. Here, we describe how the necessary, due to various reasons from sensitivity of the data to research embargo, restrictions to openness have been defined and implemented on the policy level and in technical data infrastructure solutions. We hope to contribute to an improvement of data management in the international biodiversity data infrastructures.

In Finland, the law prohibits public authorities from distributing occurrence data if this causes increased threat to endangered species. However, neither the definition of ‘endangered species’ nor guidelines for the evaluation of potential risk by openness of data are formulated. To enable mobilisation of datasets containing information on endangered species, FinBIF convened a task force commissioned to set rules on data distribution, which respect the spirit of the law. The task force consisted of representatives of relevant data holding authorities and it consulted a wide group of taxon experts and the species

information community. First, a list of species, judged to be among those targeted by the spirit of the law, was created ([sensitive species data](#)). Then the rules of restriction were decided on for each of the species. Measures of restriction ranged from complete non-disclosure of data to temporal and spatial restrictions. The identified safeguards concerning the sensitive data management in all use cases led us to create a series of innovative solutions

Researchers often wish to restrict the openness of data they have gathered for research purposes. These restrictions include embargo periods, limitations on the precision of data and controls on how the data is used. In many cases, however, researchers are willing to allow unrestricted official use of their data in certain cases such as for conservation management or land use planning. In these cases they will often allow storage and restricted use of exact data without an embargo. The same may be true for other data owners, such as nongovernmental organisations (NGOs) or private citizens. To support restrictions to openness, while simultaneously securing mobilisation of valuable datasets, FinBIF applies data sharing contracts including, as a rule, a precondition to share the original data with the public authorities for official use under the Creative Commons 4.0 BY -licence ([CC 4.0 BY](#)).

The technical solution to enabling the rather complex data policy is that FinBIF stores the collated data in two separate data warehouses: a public one for the distribution of fully open data and temporally and spatially coarsened sensitive data, alongside another containing all data but with restricted access to authorised users. In addition, to allow case-by-case release of restricted data, FinBIF has developed a data request function (Fig. 1). When users of the open data retrieve a dataset using, e.g., taxonomic and spatial filtering, they receive a search result stating whether there are restricted data available based on the filters used. In these cases a user can issue a data request, automatically distributed to all owners of data contained in the collated data batch.

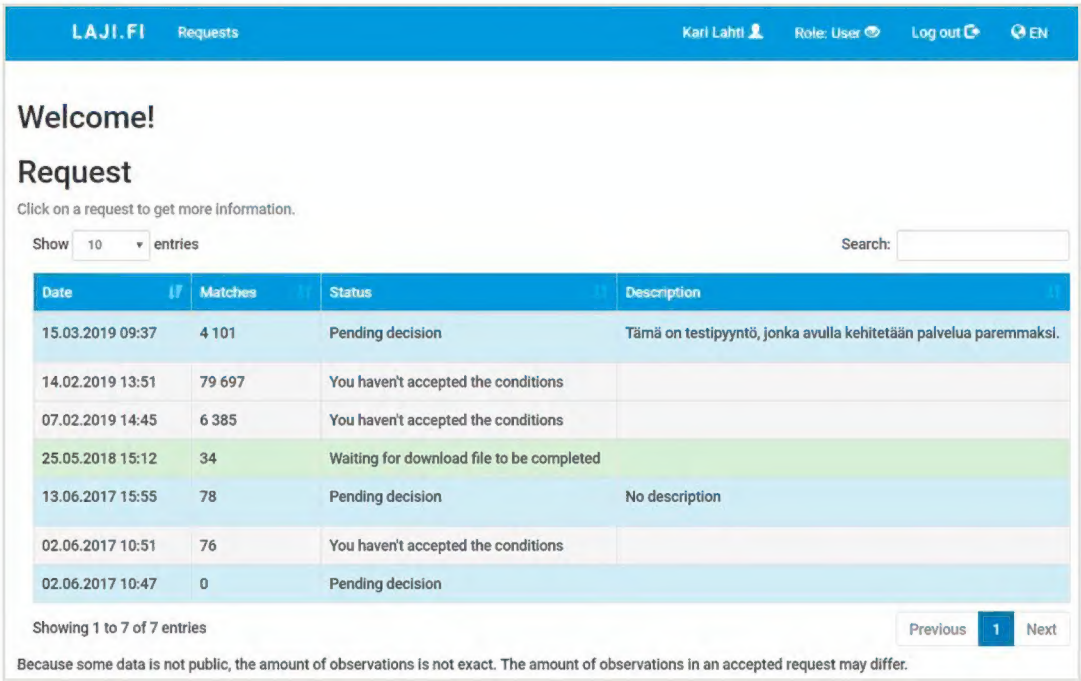


Figure 1. Restricted Data Request: service user interface for managing requesters' "ready-for-download" data or "pending issues" at FinBIF.

Agreeing on the principles about how to apply restrictions to data openness and how to define authoritative use, has not been easy given the lack of precedents. It has required thorough and inclusive consultation with both state administration, conservation practitioners, scientific specialists and lawyers. The two main cultural constraints to overcome have been (1) embracing the FAIR principles of *truly* ‘as open as possible’ and *only* ‘as closed as (absolutely) necessary’ (European Commission 2016); and, perhaps surprisingly, (2) figuring out novel ways to work across different state administrative sectors to share data.

Keywords

biodiversity informatics, sensitive data, restricted data, open data, FinBIF, GBIF, data use, FAIR

Presenting author

Kari M Lahti

Presented at

Biodiversity_Next 2019

References

- European Commission (2016) Guidelines on FAIR data management. Horizon 2020, version 3.0.
- Wilkinson MD (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3: 160018.